MOHAN LIU

https://mohanliu.github.io/
+1 847 644 9861
mohan@u.northwestern.edu
Qualify for OPT Visa

TECHNICAL SKILLS

Data Science: Data Mining (Pandas), Data Visualization (Plotly, Bokeh, Google Charts), Machine and Deep Learning, Image Processing (OpenCV), Natural Language Processing (NLTK), Time Series Forecasting (Prophet)

Machine Learning Algorithms: Support Vector Machine, Random Forest (Scikit-learn), Gradient Boosting Decision Tree (LightGBM), Neural Networks (TensorFlow, Keras, PyTorch), Latent Semantic Analysis, ARIMA

Data Engineering: SQL (MySQL, PostgreSQL), Apache Spark (PySpark), RESTful API Implementation (Django), GUI Programming (TKinter), Web Scrapping (BeautifulSoup, Selenium), Web Application (Flask, Dash)

Languages and Platforms: Python, C/C++, Git, HTML, CSS, JavaScript, Google Cloud, Docker, AWS, Heroku

EDUCATION

Ph.D. | Materials Science and Engineering — Aug. 2019
- Northwestern University (NU), Evanston, IL

B.S. | Physics — July 2013
- Nanjing University (NJU), Nanjing, China

EXPERIENCE

Research Assistant | Northwestern University — Sep. 2013 – Aug. 2019
- Managed and maintained a MySQL database that stores computed physical and chemical properties for more than 600,000 materials (around 200 GB); created a Docker image to provide the community with an easy access to our database using python API based on Django
- Implemented a RESTful API for our database using Django rest-framework that allows stateless data delivery with features including pagination, sorting, filtering and authentication; designed a python API wrapper, available in PyPI, aiming to provide a straight-forward and pythonic querying method
- Trained regression machine learning models (LASSO, SVR, random-forest and boosting decision tree) to predict materials band gaps with around 20% relative RMSE using elemental-property-based attributes
- Investigated interplay of compositional and local atomic ordering on adsorption at a bimetallic surface alloy to search for optimal catalyst with high activity and low cost
- Generated computed materials data for greater than 500 different crystal structures and applied linear regression models to predict catalytic activity at alloyed surfaces with less than 5% relative RMSE

Fellow | The Data Incubator — June 2019 – Aug. 2019
- Explored the Divvy bike sharing system in Chicago and deployed an end-to-end product to predict daily bike demand at each Divvy station on a future date
- Applied time series forecasting using historical bike-trip data considering stationarity, seasonality, special events and historical Chicago weather; decreased the error by 50% compared with baseline model
- Developed a Live Station Status Monitor to visualize the trend of available bikes/docks at each station over the past week
- Built an ETL pipeline including extracting real-time data of Divvy bike station status using cron job scheduler, transforming JSON format into PostgreSQL and loading the data into Google Cloud database

PUBLICATIONS & PRESENTATIONS

Interface and Heterostructure Design in Polyelemental Nanoparticles | Science — 2019
High-index Facet Nanoparticle-shape Regulation by Dealloying | Science — 2019
High-throughput Hybrid-functional DFT Investigations on Materials Properties | APS — 2018